

# A 5-50 Gb/s Quarter Rate Transmitter with a 4-Tap Multiple-MUX based FFE in 65 nm CMOS

Xuqiang Zheng<sup>\*†</sup>, Fangxu Lv<sup>\*</sup>, Feng Zhao<sup>†</sup>, Chun Zhang<sup>\*</sup>, Shigang Yue<sup>†</sup>, Ziqiang Wang<sup>\*</sup>, Fule Li<sup>\*</sup>, Zhihua Wang<sup>\*</sup>

<sup>\*</sup>Institute of Microelectronics, Tsinghua University, Beijing, 10084, China

<sup>†</sup>School of Computer Science, University of Lincoln, Lincoln, LN6 7TS, UK

Email: chunzhang@tsinghua.edu.cn, syue@lincoln.ac.uk

**Abstract**—This paper presents a 5-50 Gb/s quarter-rate transmitter with a 4-tap feed-forward equalization (FFE) based on multiple-multiplexer (MUX). A bandwidth enhanced 4:1 MUX with the capability of eliminating charge-sharing effect is proposed to increase the maximum operating speed. To produce the quarter-rate parallel data streams with appropriate delays, a compact latch array associated with an interleaved-retiming technique is designed. Implemented in 65 nm CMOS technology, the transmitter occupying an area of 0.6 mm<sup>2</sup> achieves a maximum data rate of 50 Gb/s with an energy efficiency of 3.1 pJ/bit.

## I. INTRODUCTION

The continuously increasing bandwidth for data communication has pushed wire-line connections towards data-rates of 50 Gb/s or beyond [1], [2]. The study group of IEEE P802.3bs has approved a 400 Gb/s standard to quadruple the backbone bandwidth of the existing 100 Gb/s Ethernet, where the most possible solutions are 16×25 Gb/s and 8×50 Gb/s [3]. Meanwhile, high-speed multi-lane connections for storage networks and embedded systems also boost a rapid growth. For example, InfiniBand has announced its 600 Gb/s HDR (in the roadmap), which adopts an architecture of 12×50 Gb/s. As one of the most important components in these serial links, the transmitter needs to produce precise timing information for correct data transmission and provide appropriate compensating abilities to cancel the channel dispersion. Moreover, the tight timing budget and high energy-efficiency requirements make the design task even more challenging. According to [3]–[5], the challenges in the transmitter around 50 Gb/s mainly concentrate on the final-stage serialization and multi-tap equalization. For the serialization, quarter-rate architecture is a promising solution because it relaxes the critical path timing margin to 3 unit interval (UI), halves the maximum clock speed, and saves considerable power [2], [4]. However, it is difficult to design the 4:1 multiplexer (MUX) in such an architecture, due to its large self-loading drain capacitance. For the multi-tap equalization, a short UI, e.g. only 20 ps for 50 Gb/s, makes it difficult to generate accurate UI-spaced serial sequences. Although feed-forward equalizations (FFE) based on analog delay line implemented in LC-cells [3] and CML-buffers [6] have been reported, they suffer from either a penalty of large area occupation or a cost of huge power consumption. Another drawback of these techniques is that the delay is susceptible to PVT variations, power fluctuation, and

substrate noise. Additionally, the operating range is limited by the delay line's adjusting range.

To address these issues, this paper presents a quarter-rate transmitter with 4-tap FFE in 65 nm CMOS process, where the UI-spaced serial data are produced by four parallel 4:1 MUXs. This scheme brings in several benefits, including compact layout implementation, accurate 1UI-delay generation, and a wide operating range. To mitigate the inherent large self-loading capacitance for quarter-rate serialization, an enhanced 4:1 MUX based on data selection is proposed, which not only improves its maximum operating speed but also effectively reduces the charge-sharing effect. In addition, a compact latch-array with an interleaved-retiming technique is adopted to produce the 16 paths of quarter-rate data streams, where the retiming clocks for both the latch array and the 4:1 MUX are implemented in power-efficiency CMOS circuits.

## II. TRANSMITTER ARCHITECTURE

Fig. 1 illustrates the block diagram of the transmitter, where a multiple-MUX based 4-tap FFE is used to compensate for the channel loss. In each tap path, a final 4:1 MUX is applied to satisfy the stringent timing requirement against PVT variations. The desired 16 paths of quarter-rate data with proper delays are produced by a compact latch array via interleaved-latching the 4-bit parallel input data. As shown at the bottom of Fig. 1, a clock path consisting of a clock conditioner and a multi-clock generator (MCG) is employed to generate various clocks. Specifically, the full swing I, Q clocks (CK0\_D/CK180\_D and CK90\_D/CK270\_D) are converted from the outputs of the CML divider, which is driven by a duty cycle clock conditioner. These clocks are further applied to four driving buffers and four pseudo-AND2s to produce 50% duty cycle clocks for the latch array and 25% duty cycle clocks for the 4:1 MUXs, respectively. Additionally, an on-chip PRBS7 generator is integrated to facilitate the performance evaluation.

Fig. 2 shows the timing diagram for the 4:1 serialization in the main tap path. The other paths are similar but with different delays. As shown in Fig. 2, the parallel quarter-rate input data of D0<n>, D1<n>, D2<n>, and D3<n> are finally relatched by PH90, PH180, PH270, and PH0 to generate the UI-spaced data streams of DMAIN0<n>, DMAIN1<n>,

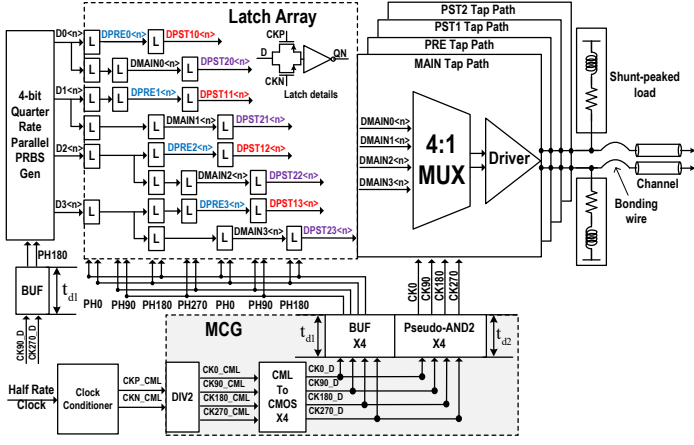


Fig. 1. Transmitter architecture.

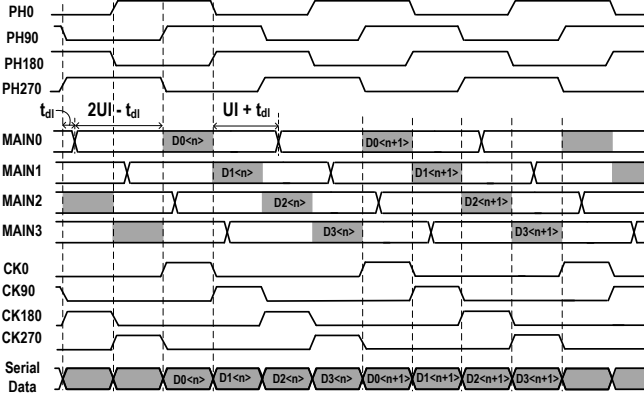


Fig. 2. Timing diagram for serialization (main tap).

DMAIN2 $\langle n \rangle$ , and DMAIN3 $\langle n \rangle$ , respectively. These prepared UI-spaced data are then sequentially selected by the 25% duty cycle clocks of CK0, CK90, CK180, and CK270 to form the serial sequence. Clearly, the margin for data selection in the 4:1 MUX is increased from 1 UI (2:1 MUX) to 3 UI, which greatly relaxes the timing constraints. To guarantee sufficient setup time and hold time for the MUX, a matched delay of  $t_{d1}$  and  $t_{d2}$  (see Fig. 1) is required. Fortunately, this can be resolved by the buffer-like pseudo-AND2, which will be detailed in Section III. Moreover, compact dynamic latches consisting of four transistors, shown as latch details in Fig. 1, are employed to further optimize the timing margin and save the power consumption. Note that the dynamic latch needs high-speed complementary clocks to frequently update the charge stored on the inverter gate capacitance to keep the data valid. This requirement is easily satisfied in this design because of the multi-GHz quarter-rate operating speed and differential clock driving scheme. On the other hand, the doubled self-drain capacitance in the 4:1 MUX significantly reduces the bandwidth of the MUX, which is the key factor that constrains its applications around the speed of 50 Gb/s. Consequently, bandwidth extending techniques for the 4:1 MUX are highly demanded.

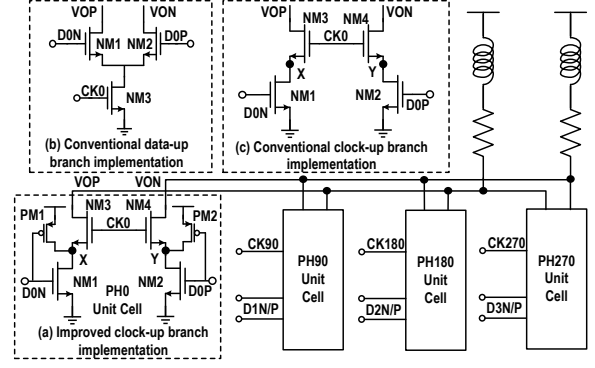


Fig. 3. The proposed 4:1 MUX.

### III. CIRCUIT DESIGN

#### A. Enhanced 4:1 MUX

Fig. 3 describes the proposed high-speed 4:1 MUX, which is composed of shunt-peaked loads and four identical pull-down unit cells. These unit cells are activated sequentially by the UI-spaced pulses (i.e., CK0, CK90, CK180, and CK270) to combine the four quarter-rate data into one serial sequence. Each cell in the MUX is implemented in a pseudo-differential structure as depicted in Fig. 3(a), eliminating the current source transistor to avoid stacked devices in critical path. Fig. 3(b) presents the traditional MUX realization described in [7], where the output can be corrupted by the data transitions on other branches through the forward coupling path from data input to the output when the MUX is performing data selection on one branch. To alleviate this problem, a clock-up MUX (Fig. 3(c)) is developed where the feed-through path is eliminated by the top clock transistor pairs [6], [8]. However, this conventional clock-up MUX suffers from severe charge-sharing effect between the outputs VOP/VON and junction nodes X/Y in forms of causing glitches on two consecutive bits at high level or slowing down the rising edges for low-to-high transitions. Specifically, if the upcoming data D0P/D0N are logic high/low, node Y is pre-discharged to ground through NM2, which helps to speed up the falling edge. On the other hand, the voltage of node X depends on previous transmitted data. In case that the previous D0N is logic low, node X should have been charged to an allowed maximum value ( $VDD - V_{THN}$ ) during the selection-enabled period (high pulse duration of CK0), which should maintain to the present instant since NM1 has always been in cut-off state. Therefore, this will not cause prominent charge extraction effect as node X has already been charged to the allowed maximum value by the previous transmitted bit. Conversely, if the previous D0N is logic high, node X should keep the ground voltage that is pulled down during the portion after data selection is disabled in previous bit period (the portion of  $UI + t_{dl}$  in Fig. 2). When the high pulse of CK0 arrives, the capacitance at node X will extract charge from the output, thus causing a remarkable glitch for two consecutive output bits at high level or slowing down the rising edge for a low-to-high transition, as

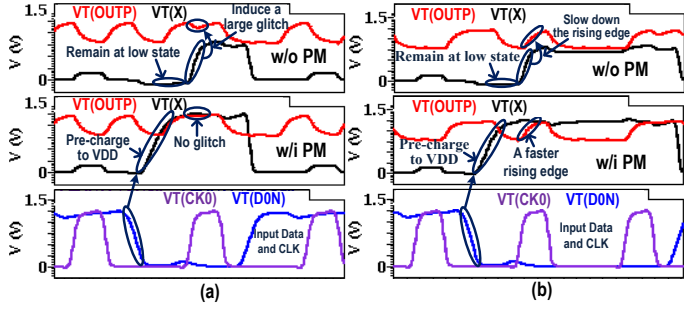


Fig. 4. Effects of the introduced PM on (a) high-level glitches and (b) edge-transitions.

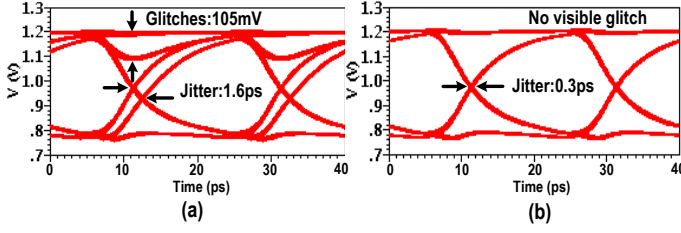


Fig. 5. Simulated eye-diagrams of 4:1 MUX. (a) Without PM and (b) with PM.

shown in the top row of Figs. 4(a) and (b), respectively. Fig. 5(a) presents the simulated eye-diagram of the multiplexed serial data, which displays apparent rising edge inter-symbol interference (ISI) and prominent glitches at high level.

To overcome these impediments, pre-charge transistors P-M1/PM2 with a small dimension connecting to nodes X/Y are added, as illustrated in Fig. 3(a). Taking the PH0 branch as an example, the operation of the proposed pull-down circuitry is as follows. When input data arrive, depending on D0N/D0P, nodes X/Y are either discharged to VSS through NM1/NM2 or pre-charged to VDD by PM1/PM2. This makes nodes X/Y always in desired states, which are coincident with the upcoming output signal levels. Then, as CK0 goes high, NM3/NM4 are turned on to send D0N/D0P to the MUX's output. After a period of 1 UI, the pull-down path is switched off by the falling edge of CK0 and the voltage level of nodes X/Y stays unchanged until the next input data come. The main feature of this novel 4:1 MUX is its ability of eliminating the charge-sharing effect caused by parasitic capacitances at nodes X/Y. This advantage brings in several benefits. First, the deterministic jitter and glitches caused by charge extraction can be significantly mitigated (see the middle row in Figs. 4(a) and (b)). The simulated eye-diagram in Fig. 5(b) indicate that the ISI caused by charge-sharing is reduced from 1.6 ps to 0.3 ps and the voltage glitches are mostly removed. Furthermore, the glitch elimination effectively improves the noise margin that allows a lower output swing to save power. Second, the elimination of the charge-sharing effect makes the capacitances at nodes X/Y less significant. Thus, large-size NM1/NM2 can be used to improve the discharging capabilities. Note that the output swing is determined by the proportion of resistive load and equivalent resistance of stacked NM1/NM3 (NM2/NM4).

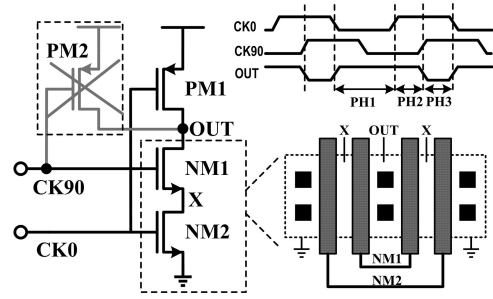


Fig. 6. Pseudo-NAND2 implementation.

For a fixed minimum output swing, the big size of NM1/NM2 indicates that NM3/NM4's size can be reduced. The small size of NM3/NM4 can effectively decrease the self-loading drain capacitance of unit cells. Consequently, the bandwidth of the overall 4:1 MUX can be improved. Third, the added PM1/PM2 provide another path through NM3/NM4 to help to pull-up the output, which can accelerate the rising transitions.

### B. 25% duty cycle clock

As the final retiming stage, the MUX performance largely relies on the retiming phases (CK0-90-180-270) depicted in Fig. 2, since any timing deviation will be converted into final output jitter directly. This necessitates the following two desirable properties: i) the high pulse width for each phase should be an accurate UI, and ii) the spacing between any two adjacent phases should be the same, which equals 1 UI. To enable a high-speed operation, these pulses are usually created by NOR/AND of two 50% duty-cycle half-rate clocks with 90° phase shifting. Since serial NMOS transistors are much faster than serial PMOS transistors, NAND2 associated with a driving inverter could be a better choice. Fig. 6 presents the designed pseudo high-speed NAND2 gate implementation. In contrary to conventional topology, the pull-up transistor PM2 is eliminated. In principle, at the beginning of PH1, node OUT is pulled up to VDD by PM1, which can be held during PH2 since NM1 is still in closed state. In PH3, both NM1 and NM2 are turned on to generate the retiming pulse. With this approach, the output capacitance can be reduced, which helps to achieve a high-speed operation. Additionally, the similar circuit realization of pseudo-AND and BUF mitigates the delay mismatching issue between  $t_{d1}$  and  $t_{d2}$  mentioned in Section II, which helps to meet the stringent timing constraints against PVT variations. It is worth noting that charge-sharing effect caused by junction capacitance and parasitic capacitance at node X still exists. In particular, when CK0 goes down to trigger PM1 to charge the output node, node X will extract charge through NM1 since CK90 is still remaining at high state. To alleviate this effect, an abutment layout approach with minimum gate spacing (see Fig. 6) is exploited to reduce the parasitic capacitance at node X. As illustrated in Fig. 6, the big serial transistors are divided into several small serial transistors, and every two small ones are connected in parallel, sharing a common drain region to reduce the junction area.

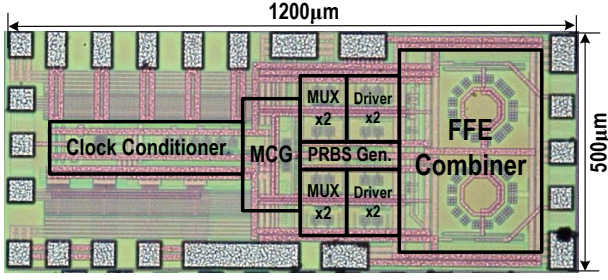


Fig. 7. Chip micrograph.

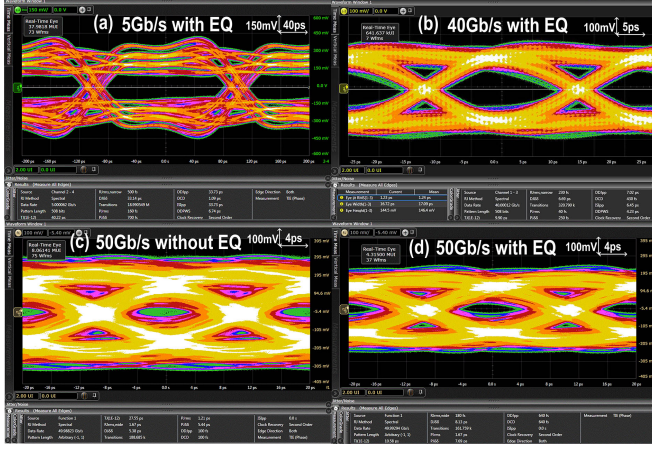


Fig. 8. Measured output eye-diagrams.

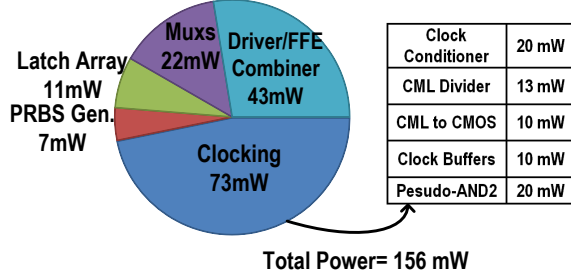


Fig. 9. Power breakdown at 50 Gb/s.

#### IV. MEASUREMENT RESULTS

Fig. 7 presents the chip micrograph of the transmitter fabricated in 65nm CMOS technology, which occupies an area of 0.6 mm<sup>2</sup> and consumes 156 mW from a 1.2 V supply. Figs. 8(a) and (b) depict the over-equalized and properly-compensated eye-diagrams at 5 Gb/s and 40 Gb/s, respectively. Particularly, at 40 Gb/s, the total jitter is 9.90 ps and the vertical eye height is 146 mV. Figs. 8(c) and (d) give the output eye-diagrams at 50 Gb/s before and after applying the 4-tap FFE. Obviously, the FFE opens the completely closed eye. The eye height and total jitter are equalized to 50 mV and 10.58 ps, respectively. It can also be seen that a wide operating range from 5 Gb/s to 50 Gb/s is achieved, which is mainly attributed to the multiple-MUX based FFE implementation. Fig. 9 shows the power breakdown of the presented transmitter, where the proposed 4:1 Muxs only consumes 22 mW.

TABLE I  
PERFORMANCE SUMMARY AND COMPARISON.

Reference	This work	[1]	[2]	[3]
Technology (nm)	65	65	14	65
Supply (V)	1.2	1.2	N/A	1.2
Data Rate (Gb/s)	5-50	60	16-40	50-64
Chip Area (mm <sup>2</sup> )	1.2 × 0.5	2.1 × 1.0	0.215 × 0.13	1.2 × 1.0
FFE	4-tap	N/A	4-tap	4-tap
1UI-delay Gen.	Multi-MUX	N/A	Multi-MUX	LC-delay
MUX Type	4:1	2:1	4:1	4:1
Data Jitter RJ (ps <sub>rms</sub> )	0.23@40Gb/s 0.18@50Gb/s	1.08@30Gb/s	0.33@28Gb/s 0.51@40Gb/s	N/A
Data Jitter (ps) TJ (BER=1e-12)	9.90@40Gb/s 10.58@50Gb/s	N/A	10.72@28Gb/s 12.89@40Gb/s	N/A
Power (mW)	156	450	518	199
Energy Efficiency (pJ/bit)	3.1	7.5	12.9	3.1

Table I summarizes the chip performance compared with other transmitters operating at similar data rates. The results indicate that this design achieves better jitter performance and power efficiency, even in comparison with the LC-delay based FFE, mainly because of the proposed high-speed 4:1 MUX and the compact interleaved-latching scheme.

#### V. CONCLUSION

The quarter-rate transmitter with 4-tap FFE is implemented in 65 nm CMOS process. The integration of a bandwidth enhanced 4:1 MUX and an interleaved-retiming latch array makes the transmitter possess good properties of both low power consumption (3.1 pJ/bit) and small area occupation (1.2 × 0.5mm<sup>2</sup>), while supporting a wide operating range of 5-50 Gb/s. Additionally, the developed transmitter exhibits a low total jitter of 10.58 ps after a 12 dB loss channel at 50 Gb/s.

#### ACKNOWLEDGMENT

This work was supported by grants from National Natural Science Foundation of China (NSFC), no. 61371011, and European FP7-LIVECODE, no. 295151.

#### REFERENCES

- [1] P. C. Chiang *et al.*, "60Gb/s NRZ and PAM4 transmitters for 400GbE in 65nm CMOS link," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2014, pp. 148–149.
- [2] J. Kim *et al.*, "A 16-to-40Gb/s quarter-rate NRZ/PAM4 dual-mode transmitter in 14nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2015, pp. 60–61.
- [3] M. S. Chen and C. K. K. Yang, "A 50-64 Gb/s serializing transmitter with a 4-tap, LC-ladder-filter-based FFE in 65 nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 50, no. 4, pp. 1903–1916, Apr. 2015.
- [4] A. A. Hafez *et al.*, "A 32-to-48Gb/s serializing transmitter using multi-phase sampling in 65nm CMOS," in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2013, pp. 38–39.
- [5] R. Navid *et al.*, "A 40 Gb/s serial link transceiver in 28 nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 50, no. 4, pp. 814–827, Dec. 2015.
- [6] B. Raghavan *et al.*, "A sub-2 W 39.8-44.6 Gb/s transmitter and receiver chipset with SFI-5.2 interface in 40 nm CMOS," *IEEE J. Solid-State Circuits*, vol. 48, no. 12, pp. 3219–3228, Dec. 2013.
- [7] H. Wang and J. Lee, "A 21-Gb/s 87-mW transceiver with FFE/DFE/Analog equalizer in 65-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 45, no. 4, pp. 909–919, Apr. 2010.
- [8] D. Cui *et al.*, "A dual-channel 23-Gbps CMOS transmitter/receiver chipset for 40-Gbps RZ-DQPSK and CS-RZ-DQPSK optical transmission," *IEEE J. Solid-State Circuits*, vol. 47, no. 12, pp. 3249–3260, Dec. 2012.